The Periodic QR Algorithm is a Disguised QR Algorithm

Daniel Kressner [1]

Institut für Mathematik MA 4-5, TU Berlin, Str. des 17. Juni 136, D-10623 Berlin, FRG.

Abstract

The periodic QR algorithm is a strongly backward stable method for computing the eigenvalues of products of matrices, or equivalently for computing the eigenvalues of block cyclic matrices. The main purpose of this paper is to show that this algorithm is numerically equivalent to the standard QR algorithm. It will be demonstrated how this connection may be used to develop a better understanding of the periodic QR algorithm.

Key words: QR algorithm, Block cyclic matrices, Matrix products
2000 MSC: 65F15

1 Introduction

In this paper we consider the eigenvalue computation of a matrix $\mathcal{A} \in \mathbb{R}^{pn \times pn}$ having the form

$$\mathcal{A} = \begin{bmatrix} 0 & & & A^{(p)} \\ A^{(1)} & \ddots & & \\ & \ddots & \ddots & \\ & & A^{(p-1)} & 0 \end{bmatrix}, \tag{1}$$

Email address: kressner@math.tu-berlin.de (Daniel Kressner).

[1] Supported by the DFG Research Center "Mathematics for key technologies" (FZT 86) in Berlin and a Marie Curie fellowship in the frame of the Control Training Site (MCFI-2001-00403).

where $A^{(k)} \in \mathbb{R}^{n\times n}$ for $k = 1, \ldots, p$. Matrices with this block cyclic structure naturally arise in applications such as periodic systems [1], queuing network models [2,3] and multiple shooting methods [4,5].

The eigenvalues of $\mathcal{A}$ are the p-th roots of the eigenvalues of the matrix product $\Pi_{\mathcal{A}} = A^{(p)}A^{(p-1)} \cdots A^{(1)}$. Unfortunately, computing the eigenvalues of the explicitely formed matrix $\Pi_{\mathcal{A}}$ is not numerically backward stable. Unless the factors $A^{(k)}$ have very low condition numbers or the eigenvalues obey an exponential splitting [6] small perturbations in the matrix $\Pi_{\mathcal{A}}$ may correspond to large backward errors in its factors. On the other hand, applying a backward stable method as the QR algorithm [7, Sec. 2.3] to $\mathcal{A}$ requires $O(p^3n^3)$ floating point operations (flops). Furthermore, it is not clear whether small backward errors in the matrix $\mathcal{A}$ can be related to small backward errors in its nonzero block entries only, which would be necessary to guarantee that the method is backward stable in a strong sense [8].

In this paper we show that the QR algorithm applied to a shuffled version of the matrix $\mathcal{A}$ completely preserves its structure and is thus not only strongly backward stable but also requires as few as $O(pn^3)$ flops. Relating it back to the block entries of the unshuffled matrix it turns out that this algorithm is numerically equivalent to the so called periodic QR algorithm [9–11]. By an analysis of the shift transmission mechanism we demonstrate how this connection can be used to generalize theoretical and practical considerations for the QR algorithm to the periodic QR algorithm.

2 The Perfect Shuffle

We will make use of a certain permutation, the perfect shuffle. In this section this permutation and some of its basic properties are reviewed. Let $z = [\ z^{(1)}, z^{(2)}, \ldots, z^{(p)}\ ]$ where $z^{(k)}$ is a row vector of length n. Imagine that each $z^{(k)}$ represents a deck of n cards. A perfect shuffle stacks exactly one card from each deck, rotationally until all decks are exhausted. The row vector that corresponds to the shuffled deck is given by

$$\tilde{z} = [\ z_1^{(1)}, z_1^{(2)}, \ldots, z_1^{(p)}, z_2^{(1)}, \ldots, z_2^{(p)}, \ldots, z_n^{(1)}, \ldots, z_n^{(p)}\ ].$$

There is a unique permutation matrix $P \in \mathbb{R}^{pn \times pn}$ such that $\tilde{z} = zP$. Applying this permutation to $\mathcal{A}$ turns it into an $n \times n$ block matrix with cyclic blocks,

$$\tilde{\mathcal{A}} := P^T \mathcal{A} P = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix}, \quad A_{ij} := \begin{bmatrix} 0 & & & a_{ij}^{(p)} \\ a_{ij}^{(1)} & \ddots & & \\ & \ddots & \ddots & \\ & & a_{ij}^{(p-1)} & 0 \end{bmatrix}. \qquad (2)$$

Any matrix of the form (2) will be called *cyclic block matrix*. Similarly, applying P to a block diagonal matrix $\mathcal{D}$ yields an $n \times n$ block matrix with $p \times p$ diagonal matrices as entries. We refer to any matrix of the latter form as *diagonal block matrix*. A class of equivalence transformations that preserves cyclic block structures and will be used here is described by the following straightforward lemma.

Lemma 1 *Let $\tilde{\mathcal{A}}$ be a cyclic block matrix and let $\tilde{\mathcal{D}}$ be an invertible diagonal block matrix. Then $\tilde{\mathcal{D}}^{-1}\tilde{\mathcal{A}}\tilde{\mathcal{D}}$ is again a cyclic block matrix.*

3 Reduction to Hessenberg Form

Reducing a general matrix $A \in \mathbb{R}^{m \times m}$ to Hessenberg form is a preliminary step in the QR algorithm in order to reduce its computational complexity. Such a reduction is usually based on Householder transformations which for any real vector x of length m are defined as

$$U_j(x) := I - 2\frac{u_j(x)u_j(x)^T}{u_j(x)^T u_j(x)}, \quad u_j(x) := N_j x - \operatorname{sign}(e_j^T x)\|N_j x\|_2 e_j, \qquad (3)$$

where $N_j := \begin{bmatrix} 0 & 0 \\ 0 & I_{m-j+1} \end{bmatrix}$ and e_j denotes the j-th unit vector of length m. Then the last $m - j$ elements of $U_j(x)^T x$ are zero. Using this definition we give a concise description of the Hessenberg reduction algorithm.

Algorithm 1 [7, Alg. 2.2] *Given a general matrix $A \in \mathbb{R}^{m \times m}$ this algorithm computes an orthogonal matrix Q such that $H = Q^T A Q$ is in upper Hessenberg form. The matrix A is overwritten by H.*

$Q \leftarrow I_m$
FOR $j \leftarrow 1, \ldots, m-2$
  $Q \leftarrow QU_{j+1}(Ae_j)$
  $A \leftarrow U_{j+1}(Ae_j)^T A U_{j+1}(Ae_j)$
END FOR

Theorem 2 *If Algorithm 1 is applied to a cyclic block matrix $\tilde{\mathcal{A}} \in \mathbb{R}^{np\times np}$ then an orthogonal diagonal block matrix $\tilde{\mathcal{Q}}$ and a cyclic block matrix $\tilde{\mathcal{Q}}^T\tilde{\mathcal{A}}\tilde{\mathcal{Q}}$ in upper Hessenberg form are returned.*

PROOF. Assume that after $(j-1)$ loops of Algorithm 1 the matrix $\tilde{\mathcal{A}}$ has been overwritten by a cyclic block matrix. Then,

$$\tilde{\mathcal{A}}e_j = y \otimes e_{k'}, \quad y = \left[\ a_{1l}^{(k)}, a_{2l}^{(k)}, \ldots, a_{nl}^{(k)}\ \right]^T, \tag{4}$$

where '$\otimes$' denotes the Kronecker product (see e.g. [7, pg. 24]), $k' = j \bmod p+1$, $k = (j-1) \bmod p+1$ and $l = (j-k)/p+1$. Since

$$u_{j+1}(\tilde{\mathcal{A}}e_j) = u_{j+1}(y \otimes e_{k'}) = \begin{cases} k < p: \ \ u_l(y) \otimes e_{k'}, \\ k = p: \ \ u_{l+1}(y) \otimes e_{k'}, \end{cases}$$

it follows that $U_{j+1}(\tilde{\mathcal{A}}e_j)$ is a diagonal block matrix. Thus, Lemma 1 shows that the j-th loop of Algorithm 1 preserves the cyclic block form of $\tilde{\mathcal{A}}$. The statement about $\tilde{\mathcal{Q}}$ is a consequence of the group property of orthogonal diagonal block matrices. $\square$

Hence, Algorithm 1 applied to $\tilde{\mathcal{A}}$ only operates on the entries $a_{ij}^{(k)}$; it should thus be possible to reformulate this algorithm in terms of operations on the factors $A^{(1)}, \ldots, A^{(p)}$. In the following we will derive such a reformulation. First note that the proof of Theorem 2 also shows that $\tilde{\mathcal{A}} \leftarrow U_{j+1}(\tilde{\mathcal{A}}e_j)^T\tilde{\mathcal{A}}U_{j+1}(\tilde{\mathcal{A}}e_j)$ is equivalent to the updates

$$\begin{cases} k < p: \ \ A^{(k+1)} \leftarrow A^{(k+1)}U_l(A^{(k)}e_l),\ A^{(k)} \leftarrow U_l(A^{(k)}e_l)^T A^{(k)}, \\ k = p: \ \ A^{(1)} \leftarrow A^{(1)}U_{l+1}(A^{(p)}e_l), \quad A^{(p)} \leftarrow U_{l+1}(A^{(p)}e_l)^T A^{(p)}, \end{cases}$$

where the quantities k, k' and l are defined as in (4). Furthermore, if we set

$$\tilde{\mathcal{Q}} = P^T \mathrm{diag}(Q^{(1)}, Q^{(2)}, \ldots, Q^{(p)})P,$$

then $\tilde{\mathcal{Q}} \leftarrow \tilde{\mathcal{Q}}U_{j+1}(\tilde{\mathcal{A}}e_j)$ equalizes $Q^{(k+1)} \leftarrow Q^{(k+1)}U_l(A^{(k)}e_l)$ for $k < p$ and $Q^{(1)} \leftarrow Q^{(1)}U_{l+1}(A^{(p)}e_l)$ for $k = p$. Altogether, we can rewrite Algorithm 1 in the following way.

Algorithm 2 *Given the matrices $A^{(1)}, \ldots, A^{(p)} \in \mathbb{R}^{n\times n}$ this algorithm computes orthogonal matrices $Q^{(1)}, \ldots, Q^{(p)}$ such that $H^{(k)} = Q^{(k+1)T}A^{(k)}Q^{(k)}$ is upper triangular for $k = 1, \ldots, p-1$ and $H^{(p)} = Q^{(1)T}A^{(p)}Q^{(p)}$ is in upper Hessenberg form. Each matrix $A^{(k)}$ is overwritten by $H^{(k)}$.*

$Q^{(1)} \leftarrow I_n, Q^{(2)} \leftarrow I_n, \ldots, Q^{(p)} \leftarrow I_n$

FOR $l \leftarrow 1, \ldots, n-1$
  FOR $k \leftarrow 1, \ldots, p-1$
    $Q^{(k+1)} \leftarrow Q^{(k+1)} U_l(A^{(k)} e_l)$
    $A^{(k+1)} \leftarrow A^{(k+1)} U_l(A^{(k)} e_l)$
    $A^{(k)} \leftarrow U_l(A^{(k)} e_l)^T A^{(k)}$
  END FOR
  $Q^{(1)} \leftarrow Q^{(1)} U_{l+1}(A^{(p)} e_l)$
  $A^{(1)} \leftarrow A^{(1)} U_{l+1}(A^{(p)} e_l)$
  $A^{(p)} \leftarrow U_{l+1}(A^{(p)} e_l)^T A^{(p)}$
END FOR

Note that this Algorithm corresponds to the reduction to periodic Hessenberg form described in [9, Pg. 5–7]. It should be emphasized that Algorithm 2 performs exactly the same operations as Algorithm 1 applied to $\tilde{\mathcal{A}}$. Hence, also in the presence of roundoff errors both algorithms produce the same result, an entity that is commonly called *numerical equivalence*.

Example 3 *If $p = 2$ and $A^{(1)} = A^{(2)T}$ then Algorithm 1 reduces the symmetric matrix $\tilde{\mathcal{A}}$ to tridiagonal form. On the other hand, Algorithm 2 returns $Q^{(2)T} A^{(1)} Q^{(1)}$ in bidiagonal form. Hence, as a special case we obtain that bidiagonal reduction [7, Alg. 3.2] applied to $A^{(1)}$ is numerically equivalent to tridiagonal reduction applied to $\tilde{\mathcal{A}}$. A similar observation has been made by Paige [12].*

4 QR Steps

Given a set of shifts $\Sigma = \{\sigma_1, \ldots, \sigma_s\} \subset \mathbb{C}$, closed under complex conjugation, the explicitly shifted QR step applied to a general matrix $A \in \mathbb{R}^{m \times m}$ computes a QR decomposition

$$(A - \sigma_1 I_m)(A - \sigma_2 I_m) \cdots (A - \sigma_s I_m) = QR$$

and performs the update $A \leftarrow Q^T A Q$. If A is in *unreduced* Hessenberg form, i.e., $a_{j+1,j} \neq 0$ for $j = 1, \ldots, n-1$, then an explicitly shifted QR step is equivalent to an implicitly shifted QR step, described by the following algorithm. Note that we let Σ contain the Wilkinson shifts, these are the eigenvalues of the bottom right $s \times s$ submatrix of A.

Algorithm 3 [7, Alg. 2.8] *Given a Hessenberg matrix $A \in \mathbb{R}^{m \times m}$ this algorithm performs an implicitly shifted QR step with s Wilkinson shifts on A and returns the corresponding orthogonal transformation matrix Q.*

Compute $\{\sigma_1, \ldots, \sigma_s\}$ as the eigenvalues of $A(m-s+1:m, m-s+1:m)$.
$x \leftarrow (A - \sigma_1 I_m)(A - \sigma_2 I_m) \cdots (A - \sigma_s I_m) e_1$

$A \leftarrow U_1(x)^T A U_1(x)$
Apply Algorithm 1 to compute an orthogonal matrix Q such that
$A \leftarrow Q^T A Q$ is in Hessenberg form.
$Q \leftarrow U_1(x) Q$

Again, cyclic block structures are preserved if s is wisely chosen.

Theorem 4 *If Algorithm 3 is applied to a cyclic block matrix $\tilde{\mathcal{A}} \in \mathbb{R}^{np \times np}$ in Hessenberg form and the number of shifts is an integer multiple of p, say $s = pt$, then the structure of $\tilde{\mathcal{A}}$ is preserved and an orthogonal diagonal block matrix $\tilde{\mathcal{Q}}$ is returned.*

PROOF. The bottom right $s \times s$ submatrix of $\tilde{\mathcal{A}}$ is a cyclic block matrix. Thus, the Wilkinson shifts can be partitioned into groups $\{\sigma_i^{(1)}, \ldots, \sigma_i^{(p)}\}$, $i = 1, \ldots, t$, where each group contains the p-th roots of some $\gamma_i \in \mathbb{C}$. Using the fact that $\Pi_{\mathcal{A}} = A^{(p)} A^{(p-1)} \cdots A^{(1)}$ is the leading $n \times n$ block of the block diagonal matrix $(P \tilde{\mathcal{A}} P^T)^p$ we obtain

$$
\begin{aligned}
x &= \prod_{i=1}^{t} \prod_{k=1}^{p} (\tilde{\mathcal{A}} - \sigma_i^{(k)} I_{np}) e_1 = \prod_{i=1}^{t} \left(\tilde{\mathcal{A}}^p - \prod_{k=1}^{p} \sigma_i^{(k)} I_{np} \right) e_1 \\
&= P^T \cdot \prod_{i=1}^{t} \left((P \tilde{\mathcal{A}} P^T)^p - \gamma_i I_{np} \right) P e_1 = \left(\prod_{i=1}^{t} (\Pi_{\mathcal{A}} - \gamma_i I_n) e_1 \right) \otimes e_1.
\end{aligned}
$$

Thus, $U_1(x)$ is block diagonal, which together with Theorem 2 concludes the proof. $\square$

The subdiagonal of $\tilde{\mathcal{A}}$ consists of the diagonals of $A^{(1)}, \ldots, A^{(p-1)}$ and the subdiagonal of $A^{(p)}$. Hence, the Hessenberg matrix $\tilde{\mathcal{A}}$ is unreduced if and only if all the triangular factors are nonsingular and the Hessenberg factor is unreduced. Similar to Hessenberg reduction the proof of Theorem 4 gives a way to rewrite Algorithm 3 in terms of operations on the factors of $\tilde{\mathcal{A}}$.

Algorithm 4 *Given nonsingular upper triangular matrices $A^{(1)}, \ldots, A^{(p-1)} \in \mathbb{R}^{n \times n}$ and an unreduced Hessenberg matrix $A^{(p)} \in \mathbb{R}^{n \times n}$ this algorithm computes orthogonal matrices $Q^{(1)}, \ldots, Q^{(p)}$ so that the updated $A^{(1)}, \ldots, A^{(p)}$ are the factors of the cyclic block matrix that would have been obtained after one QR step with pt shifts has been applied to $\tilde{\mathcal{A}}$.*

Compute $\{\gamma_1, \ldots, \gamma_t\}$ as the eigenvalues of $\Pi_A(n-t+1:n, n-t+1:n)$.
$x \leftarrow (\Pi_A - \gamma_1 I_n)(\Pi_A - \gamma_2 I_n) \cdots (\Pi_A - \gamma_t I_n) e_1$
$A^{(1)} \leftarrow A^{(1)} U_1(x), \quad A^{(p)} \leftarrow U_1(x)^T A^{(p)}$
Apply Algorithm 1 to compute orthogonal matrices $Q^{(1)}, \ldots, Q^{(p)}$ so that
$A^{(1)} \leftarrow Q^{(2)T} A^{(1)} Q^{(1)}, \ldots, A^{(p-1)} \leftarrow Q^{(p)T} A^{(p-1)} Q^{(p-1)}$ are upper

triangular and $A^{(p)} \leftarrow Q^{(1)T} A^{(p)} Q^{(p)}$ is in Hessenberg form.
$Q^{(1)} \leftarrow U_1(x) Q^{(1)}$

This algorithm is a 'Householder version' of the periodic QR step with t shifts [9, pg. 11–12].

Example 5 *This is a continuation of Example 3. If $A^{(1)}$ and $A^{(2)} = A^{(1)T}$ satisfy the assumptions of Algorithm 4 then $A^{(1)}$ is an bidiagonal matrix with nonzero diagonal and supdiagonal elements. Algorithm 3 applied to the tridiagonal matrix $\tilde{\mathcal{A}}$ performs an implicitly shifted symmetric QR step [7, Alg. 1.3] and Algorithm 4 performs a bidiagonal QR step [7, Alg. 3.4]. This shows that both QR steps are numerically equivalent.*

5 Deflation Strategies

A deflation occurs when one of the subdiagonal entries becomes sufficiently small. The usual criterion is to declare a subdiagonal entry negligible if it is small compared to the neighboring diagonal elements. This is however not a very sensible choice for matrices with zero diagonal like $\tilde{\mathcal{A}}$. Considering the action of the Householder transformations in the course of one QR step it is advisable to base the criterion on the two closest nonzero elements in the same row and column. Suitable generic criteria for $\tilde{\mathcal{A}}$ are given by

$$|a^{(p)}_{j+1,j}| \leq \epsilon(|a^{(p)}_{j,j}| + |a^{(p)}_{j+1,j+1}|), \tag{5}$$

$$|a^{(k)}_{j,j}| \leq \epsilon(|a^{(k)}_{j-1,j}| + |a^{(k)}_{j,j+1}|), \quad k = 1, \ldots, p-1, \tag{6}$$

where ϵ is a chosen tolerance and an entry on the right hand side of (6) is replaced by zero if it does not exist. Note that inequality (6) may only be satisfied if the 2-norm condition number of $A^{(k)}$ is not less than $1/(2\epsilon)$.

Situation (5) is easily handled, setting $a^{(p)}_{j+1,j}$ zero makes $\tilde{\mathcal{A}}$ block upper triangular,

$$\tilde{\mathcal{A}} = \begin{bmatrix} \tilde{\mathcal{A}}_{11} & \tilde{\mathcal{A}}_{12} \\ 0 & \tilde{\mathcal{A}}_{22} \end{bmatrix},$$

where $\tilde{\mathcal{A}}_{11} \in \mathbb{R}^{jp \times jp}$ and $\tilde{\mathcal{A}}_{11} \in \mathbb{R}^{(j-1)p \times (j-1)p}$ are cyclic block matrices. In contrast, situation (6) yields a deflation into two smaller eigenproblems which do not carry the structure of $\tilde{\mathcal{A}}$. For illustration, consider the case $p = n = 3$

and $a_{22}^{(2)} = 0$:

$$\tilde{\mathcal{A}} = \left[\begin{array}{ccc|cc||c|cc} 0 & 0 & X & 0 & 0 & X & 0 & 0 & X \\ X & 0 & 0 & X & 0 & 0 & X & 0 & 0 \\ 0 & X & 0 & 0 & X & 0 & 0 & X & 0 \\ \hline 0 & 0 & X & 0 & 0 & X & 0 & 0 & X \\ 0 & 0 & 0 & X & 0 & 0 & X & 0 & 0 \\ \hline\hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & X & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & X & 0 & 0 & X \\ 0 & 0 & 0 & 0 & 0 & 0 & X & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & X & 0 \end{array}\right].$$

Fortunately, there is an easy way to force deflations at $a_{21}^{(3)}$ and $a_{32}^{(3)}$ so that afterwards the deflation stemming from $a_{22}^{(2)}$ resides in a deflated $p \times p$ cyclic matrix and can thus be ignored. Applying an implicitly shifted QR step with p zero shifts introduces the zero $a_{21}^{(3)}$ element. An RQ step is a QR step implicitly applied to $(F^T \tilde{\mathcal{A}} F)^T$, where F is the flip matrix. Hence, an implicitly shifted RQ step with p zero shifts preserves the structure of $\tilde{\mathcal{A}}$ and gives us the zero $a_{32}^{(3)}$ element. Using the results of Section 4 it is easy to observe that this procedure is numerically equivalent to the deflation strategy presented in [9, pg. 7–9] apart from the fact that criterion (6) is based on the norm of $A^{(k)}$ in the latter strategy.

6 Transmission of Shifts

This section shall demonstrate how the numerical equivalence between periodic and standard QR may lead to a better understanding of the former algorithm. In the following we assume that $\tilde{\mathcal{A}}$ is in unreduced Hessenberg form. Let

$$\Sigma = \bigcup_{i=1}^{t} \{\sigma_i^{(1)}, \ldots, \sigma_i^{(p)}\}, \quad \left(\sigma_i^{(k)}\right)^p = \gamma_i,$$

be the set of shifts and let

$$\tilde{x} = \prod_{i=1}^{t} \prod_{k=1}^{p} (\tilde{\mathcal{A}} - \sigma_i^{(k)} I_{np}) e_1.$$

Then the proof of Theorem 4 shows that $\tilde{x} = x \otimes e_1$ where $x = \prod_{i=1}^{t}(\Pi_{\mathcal{A}} - \gamma_i I_n)$. Let us consider the *initial bulge pencil*

$$\tilde{\mathcal{B}}_0 - \lambda N := \begin{bmatrix} x_1 e_1 & A_{11} & \dots & \dots & A_{1t} \\ x_2 e_1 & A_{21} & \dots & \dots & A_{2t} \\ \vdots & 0 & \ddots & & \vdots \\ x_t e_1 & \vdots & \ddots & A_{t,t-1} & A_{tt} \\ x_{t+1} & 0 & \dots & 0 & a_{t+1,t}^{(p)} e_p^T \end{bmatrix} - \lambda N, \tag{7}$$

where the blocks A_{ij} are defined as in (2) and N is the $(tp+1) \times (tp+1)$ nilpotent Jordan block. The tip of the bulge, $x_{t+1} = a_{t+1,t}^{(p)} a_{tt}^{(p-1)} \cdots a_{tt}^{(1)}$, cannot be zero. It follows from a result by Watkins [13, Thm. 1] that the eigenvalues of $\tilde{\mathcal{B}}_0 - \lambda N$ are given by $\Sigma \cup \{\infty\}$.

If an implicit QR step as described in Algorithm 3 is applied to $\tilde{\mathcal{A}}$ then this matrix is updated by $U_1(\tilde{x})^T \tilde{\mathcal{A}} U_1(\tilde{x})$ in the first step. The update destroys the Hessenberg structure in rows $2, \dots, tp+1$ and columns $1, \dots, tp$. The corresponding submatrix, which is usually referred to as the *bulge*, can be partitioned as

$$\tilde{\mathcal{B}}_1 = \begin{bmatrix} a \otimes e_1 & \tilde{\mathcal{C}} \\ \alpha & b^T \otimes e_p^T \end{bmatrix}, \tag{8}$$

where $\tilde{\mathcal{C}} \in \mathbb{R}^{tp \times tp}$ is a cyclic block matrix and $a, b \in \mathbb{R}^t$, $\alpha \in \mathbb{R} \backslash \{0\}$. Again, the eigenvalues of the bulge pencil $\tilde{\mathcal{B}}_1 - N$ are $\Sigma \cup \{\infty\}$ [13, Thm. 2]. In each loop of the subsequent reduction to Hessenberg form the bulge is pushed one position along the subdiagonal to the south-east corner of the matrix. After the j-th loop the bulge resides in rows $j+2, \dots, tp+j+1$ and columns $j+1, \dots, tp+j$ and has the same structure and spectral properties as $\tilde{\mathcal{B}}_1$.

The numerical experiments conducted in [13] show strong evidence that bulge pencils whose non-infinite eigenvalues are very sensitive to perturbations will have a negative influence on the convergence of the QR algorithm. As the QR algorithm preserves block cyclic structures we only need to consider these perturbations that preserve the generic structure of bulge pencils. A good measure for the eigenvalue sensitivities might thus be given by the following component-wise condition number.

Lemma 6 *Let $\mathcal{C}, \mathcal{E} \in \mathbb{R}^{pt \times pt}$ be block cyclic matrices, i.e., they have the form (1). Further, let $\triangle\mathcal{C} \in \mathbb{R}^{pt \times pt}$ and $a, b, \triangle a, \triangle b \in \mathbb{R}^t$, $\alpha \in \mathbb{R} \backslash \{0\}, \triangle\alpha \in \mathbb{R}$ satisfy $|\triangle\mathcal{C}| \leq \epsilon\mathcal{E}, |\triangle a| \leq \epsilon|a|, |\triangle b| \leq \epsilon|b|, |\triangle\alpha| \leq \epsilon|\alpha|$ for some $\epsilon > 0$. Then*

for any non-infinite eigenvalue $\hat{\lambda}$ of the perturbed pencil

$$\begin{bmatrix} (a+\triangle a)\otimes e_1 & P^T(\mathcal{C}+\triangle\mathcal{C})P \\ \alpha+\triangle\alpha & (b+\triangle b)^T\otimes e_p^T \end{bmatrix} - \lambda N \tag{9}$$

there exists an eigenvalue λ of

$$\begin{bmatrix} a\otimes e_1 & P^T\mathcal{C}P \\ \alpha & b^T\otimes e_p^T \end{bmatrix} - \lambda N \tag{10}$$

so that

$$\frac{|\hat{\lambda}-\lambda|}{|\lambda|} \le \frac{|\alpha||u|^T\mathcal{E}|v| + 3|a|^T|u^{(1)}||b|^T|v^{(1)}|}{\alpha|u|^T|v|}\epsilon + O(\epsilon^2), \tag{11}$$

where $u = [\ u^{(1)T},\ldots,u^{(p)T}\ ]^T$, $v = [\ v^{(1)T},\ldots,v^{(p)T}\ ]^T$ are the left and right eigenvectors of $\mathcal{C} - 1/\alpha\cdot e_1e_p^T\otimes ab^T$ associated with the eigenvalue λ.

PROOF. By direct computation one can show that

$$\tilde{u} = \begin{bmatrix} P^Tu \\ -(e_1\otimes a)^Tu/\alpha \end{bmatrix}, \quad \tilde{v} = \begin{bmatrix} -(e_p\otimes b)^Tv/\alpha \\ P^Tv \end{bmatrix},$$

are the left and right eigenvectors of (10) associated with λ. By a result of Higham and Higham [14, Thm. 3.2] we obtain that

$$|\hat{\lambda}-\lambda|/|\lambda| \le \mathrm{cond}(\lambda)\epsilon + O(\epsilon^2)$$

with the eigenvalue condition number

$$\mathrm{cond}(\lambda) := \frac{1}{|\tilde{u}|^TN|\tilde{v}|}|\tilde{u}|^T \begin{bmatrix} |a\otimes e_1| & |P^T\mathcal{E}P| \\ |\alpha| & |b\otimes e_p|^T \end{bmatrix} |\tilde{v}|.$$

Now, inequality (11) follows from

$$\mathrm{cond}(\lambda) = \frac{|\alpha||u|^T\mathcal{E}|v| + |a^Tu^{(1)}|(|b|^T|v^{(1)}| + |b^Tv^{(1)}|) + |a|^T|u^{(1)}||b^Tv^{(1)}|}{|\alpha||u|^T|v|}.$$

□

7 Concluding Remarks

We have shown that Hessenberg reduction as well as QR iterations preserve cyclic block structures. If the factors $A^{(1)}, \ldots, A^{(p-1)}$ are sufficiently well conditioned then the complete QR algorithm is structure-preserving. Otherwise, a special deflation technique, which is not part of the standard QR, must be used. We hope that this connection may lead to a better understanding not only of the periodic QR algorithm but also of other algorithms used for analyzing and designing periodic systems [1].

The results in this paper can be generalized to pencils of the form $\lambda\tilde{\mathcal{E}} - \tilde{\mathcal{A}}$ where $\tilde{\mathcal{E}}$ is a diagonal block matrix and $\tilde{\mathcal{A}}$ a cyclic block matrix. It turns out that the periodic QZ algorithm [9,10] with t shifts is numerically equivalent to the QZ algorithm [7, Sec. 4] with $s = tp$ shifts, which preserves the structure of $\lambda\tilde{\mathcal{E}} - \tilde{\mathcal{A}}$.

Acknowledgements

The author thanks Prof. Volker Mehrmann for useful discussions.

References

[1] A. Varga, P. V. Dooren, Computational methods for periodic systems - an overview, in: Proc. of IFAC Workshop on Periodic Control Systems, Como, Italy, 2001, pp. 171–176.

[2] F. Bonhoure, Y. Dallery, W. J. Stewart, On the use of periodicity properties for the efficient numerical solution of certain Markov chains, Numer. Linear Algebra Appl. 1 (3) (1994) 265–286.

[3] W. J. Stewart, Introduction to the numerical solution of Markov chains, Princeton University Press, Princeton, NJ, 1994.

[4] U. M. Ascher, R. M. M. Mattheij, R. D. Russell, Numerical solution of boundary value problems for ordinary differential equations, Vol. 13 of Classics in Applied Mathematics, SIAM, Philadelphia, PA, 1995.

[5] K. Lust, Continuation and bifurcation analysis of periodic solutions of partial differential equations, in: Continuation methods in fluid dynamics (Aussois, 1998), Vieweg, Braunschweig, 2000, pp. 191–202.

[6] S. Oliveira, D. E. Stewart, Exponential splittings of products of matrices and accurately computing singular values of long products, in: Proceedings

of the International Workshop on Accurate Solution of Eigenvalue Problems (University Park, PA, 1998), Vol. 309:1-3, 2000, pp. 175–190.

[7] G. W. Stewart, Matrix algorithms. Vol. II, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001, eigensystems.

[8] J. R. Bunch, The weak and strong stability of algorithms in numerical linear algebra, Linear Algebra Appl. 88/89 (1987) 49–66.

[9] A. Bojanczyk, G. H. Golub, P. V. Dooren, The periodic Schur decomposition; algorithm and applications, in: Proc. SPIE Conference, Vol. 1770, 1992, pp. 31–42.

[10] J. J. Hench, A. J. Laub, Numerical solution of the discrete-time periodic Riccati equation, IEEE Trans. Automat. Control 39 (6) (1994) 1197–1210.

[11] C. F. Van Loan, A general matrix eigenvalue algorithm, SIAM J. Numer. Anal. 12 (6) (1975) 819–834.

[12] C. C. Paige, Bidiagonalization of matrices and solutions of the linear equations, SIAM J. Numer. Anal. 11 (1974) 197–209.

[13] D. S. Watkins, The transmission of shifts and shift blurring in the QR algorithm, in: Proceedings of the Fourth Conference of the International Linear Algebra Society (Rotterdam, 1994), Vol. 241/243, 1996, pp. 877–896.

[14] D. J. Higham, N. J. Higham, Structured backward error and condition of generalized eigenvalue problems, SIAM J. Matrix Anal. Appl. 20 (2) (1999) 493–512.